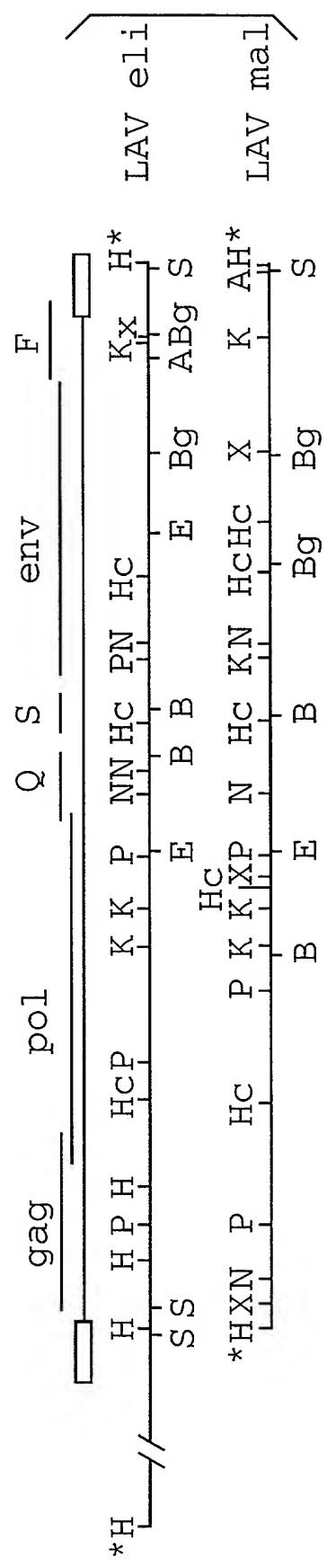
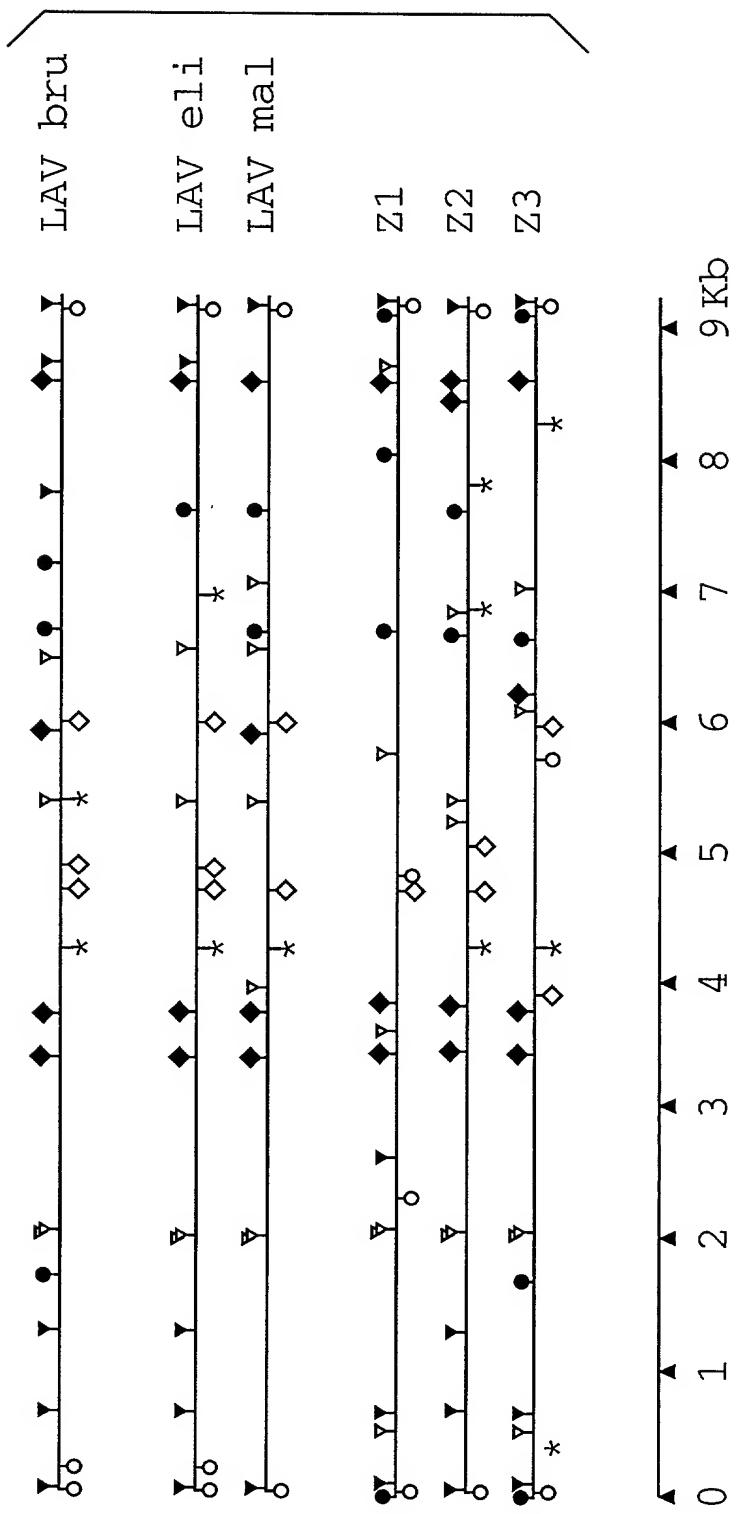


\*\*H HXN P Hc KXP Hc K N Hc KN Hc Hc X K AH\* S Bg Bg Bg Bg S LAV ma]



**FIG. 1A**



**FIG. 1B**

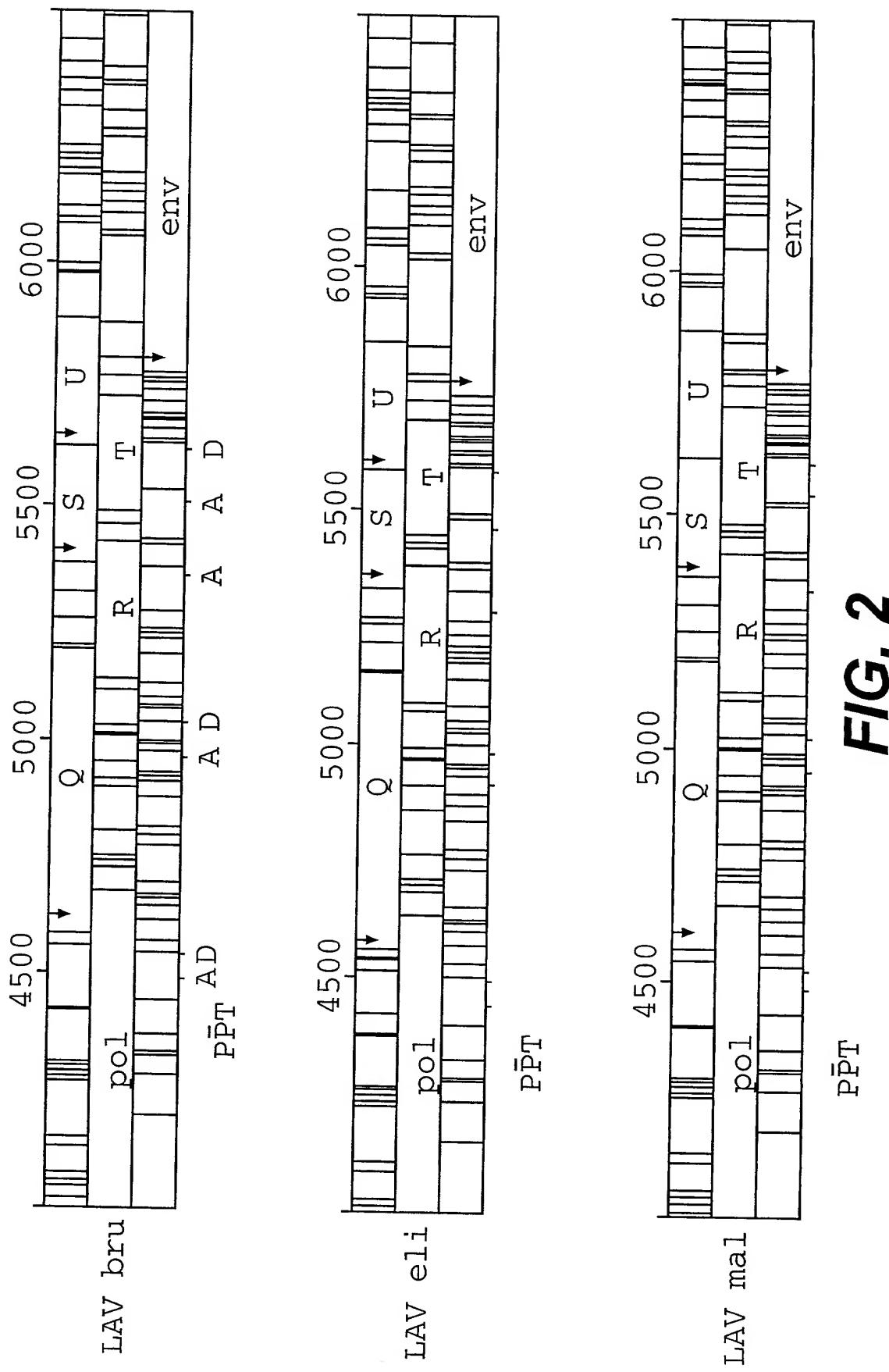


FIG. 2

GAG		10	20	30	40	50	60	70	80
LAV BRU	MGARASVLSG	GELDRWEKIR	LPGGKKKYK	LKHIVMASRE	LERFAWNPL	LETSEGCRQI	LGQLQPSLQT	GSEELRSLYN	
ARV 2		K				C			
LAV MAL		K A		R L		K	ME	ST K	IK
LAV ELI		K K		R	Y L	I		AI	T
						↓p25			
LAV BRU	TVATLYCVHQ	RIEIKDTKEA	LDKIEEQNK	SKKKAAQQAAA	-----	DTGHH	SSQVSQNYPI	VQNIQGQMVH	QAISPRTLNA
ARV 2		DV	E			---AAG N		L	
LAV MAL		DV	I	RQ T	AQQAAAAA	KN S	A	I	
LAV ELI		K G DV	E M		-----	N N	-----	L	
LAV BRU	WVKVVEEKAF	SPREVIPMFSA	LSCCGATPQDL	NTMLNTVGGH	QAAMQMLKET	INEEAAEWDK	VHPVHAGPIA	PGQMREPRGS	
ARV 2									
LAV MAL		I		M I		D	D	P	
LAV ELI		I					L		
LAV BRU	DIAGTTSTLQ	EQIGWMNNP	PIPVGTEIYKR	WILLGLNKIV	RMYSPSTILD	IRQGPKEPFR	DYVDRFYKTL	RAEQASQEVK	
ARV 2									
LAV MAL		S	D	V	V		F	T	
LAV ELI		A S	V	V	V		D		

**FIG. 3A-1**

the first time in the history of the world, the *whole* of the human race, in all its parts, has been brought together in one common cause.

LAV	BRU	NWMTETILLVQ	NANPDCCKTIL	KALGPAATLE	EMMTACQGVG	GPGHKARVIA	EAMSQVTINS-	ATIMMQRGMF	RNQRKIVKCF
ARV	2						P-	N	
LAV	MAL			G		S	A	T	KG - RI
LAV	ELI			Q		S	A	V T A	KGP I
LAV	BRU	NCGKEGHIAR	NCRAPRKKG	WKGKGKEGHQM	KDCTERQANF	LGK1WPSYKG	RPGNFLQSRP	EPTAPPFLQS	RPEPTAPPEE
ARV	2			K	R	R			
LAV	MAL			L			H		
LAV	ELI			K	R	L	R	H	
LAV	BRU	SFRSGVETT	PSQKQEPIDK	ELYPLTSIRS	LGNDPSSQ				
ARV	2	F E K							
LAV	MAL	GF E IK-	QK	A	K	QL			
LAV	ELI	GF E I -	QK	K	K				

FIG. 3A-2

CENTRAL REGION:		Q									
		10	20	30	40	50	60	70	80		
LAV	BRU	MENRWQVMIV	WQVDRMRIRT	WKSIVKHHMY	VSGKARGWMFY	RHHYESPHPR	ISSEVHPLG	DARLVITIY	GLHTGERDWH		
ARV	2			I	K	T	V	K	E		
LAV	MAL		H	K	KN	R	K	V			
LAV	ELI		K	K	NR	K	E	VR	Q	K	E
								E	K		
LAV	BRU	LGQGVSIEMR	KKRYSTQVDP	ELADQLIHY	YFDCFSSDSAI	RKALLGHIVS	PRCEYQAGHN	KVGSLOQYLAL	AALITPKKIK		
ARV	2	A	K	G	H	E	KN	I	YR	T	
LAV	MAL	H	Q	L	D	E	Q	I	D	T	A
LAV	ELI		R	G	M	E	I	D		T	A
										Q	
LAV	BRU	PPLPSVTKLIT	EDRMWNKPQKT	KGHRGSHTMN	GH						
ARV	2		K								
LAV	MAL		R								
LAV	ELI		Q	Q	R						

**FIG. 3B-1**

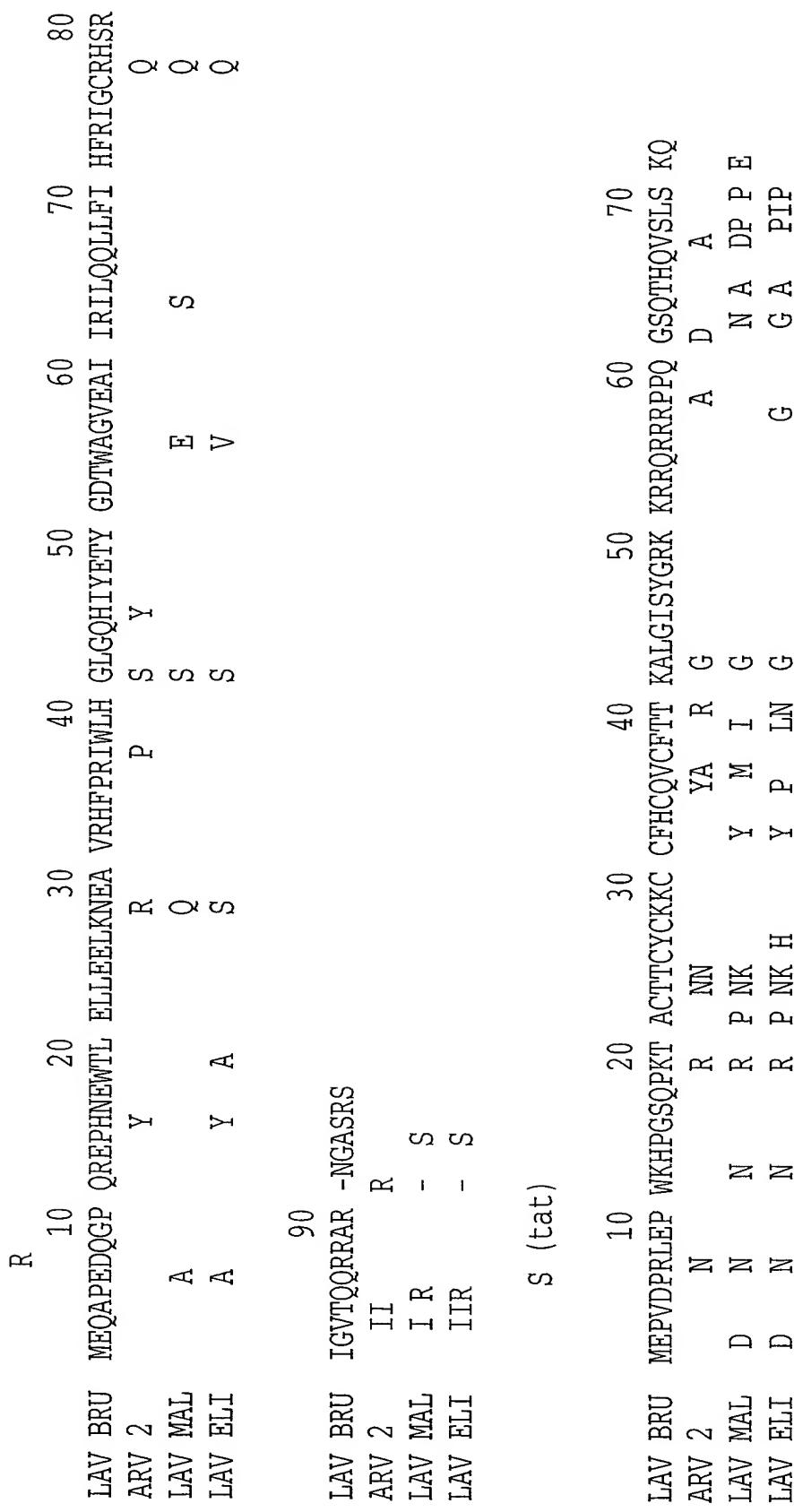


FIG. 3B-2

POL		10	20	30	40	50	60	70	80
LAV BRU	FFREDLAFLQ	GRAREFSSEQ	TRANSPTSS	EQTRANSPTTR	RELQVWGRDN	NSLSEAGADR	QGTVSFNFPQ	ITLMQRPLVT	
ARV 2						GE			
LAV MAL	N	P	P			R	G - KT	I	
LAV ELI	N	P	G L	PK		R	- P	KT	E
									A
									V
LAV BRU	IKIGQLKEA	LLDTGADDTV	LEEMSLPGRW	KPKMIGGIGG	FIKVQYDQI	LIEICGHKAI	GTVLVGPTPV	NIIGRNLITQ	
ARV 2	R					PV			
LAV MAL	VRV					P	K	I	
LAV ELI						Q			
									M
LAV BRU	IGCTLNFPIS	PIETVVPVRLK	PGMDGPVKVQ	WPLTEEKIKA	LVEICTEMEK	EGKISKIGPE	NPYNTTPVFAI	KKKDSTKWRK	
ARV 2									
LAV MAL			R		T	KD	L		
LAV ELI					T	D	R	I	
LAV BRU	LVDFRELNKR	TQDFWEVQLG	IPHPAGLKKK	KSVTVLDVGD	AYFSVPLDED	FRKYTAFTIP	SINNETPGIR	YQYNVLPQGW	
ARV 2									
LAV MAL	N								
LAV ELI									S

# FIG. 3C-1

www.3dprintables.com - 3D printer files for free

LAV	BRU	KGSPAIFQSS	MTKILEPFRK	QMPDIVIYQY	MDDLYVGSDL	EIGQHRTKIE	ELRQHLLRWG	LTPDKKHQK	EPPFLIMGYE	400
ARV	2									
LAV	MAI		T	K	E		E	K	F	
LAV	ELI			EM			K	E	F	R
LAV	BRU	LHPDKWTVQP	IVLPPEKDSWT	VNDIQKLVGK	LNWASQIYPG	IKVRQLCKLL	RGTKALTEVI	PLTSEEAEEL	AENREILKEP	480
ARV	2		M		A	K				
LAV	MAI		Q	D	E		K			
LAV	ELI		S	K	E	N	ER			
LAV	BRU	VHGVVYDPSK	DLIAEIQKQG	QGQWTYQIYQ	EPFKNLKTGK	YARTRGAHTN	DVKQLTEAVQ	KITTESVIW	GKTPPKFLPI	560
ARV	2		E	V		M		S		
LAV	MAI			QY	IKS		AQ		I	
LAV	ELI			H	M		A	R	R	R

**FIG. 3C-2**

570	580	590	600	610	620	630	640
LAV BRU	QKETWETWMT	EWQATWIE	WEFVNTPPLV	KIWyQLEKEP	IVGAETFYVD	GAASRETKLG	KAgyVTNRR
ARV 2	A	M				N	SIA
LAV MAL	A			T		N	S
LAV ELI	A				I	K	E
						N	P
650	660	670	680	690	700	710	720
LAV BRU	NQKTELQAIH	LALQDSGLEV	NIVTDSQYAL	GIQQAQPDKS	ESELVNQIE	QIJKKEKVYL	AWVPAHKGIG
ARV 2				S	I		GNEQVDKLVS
LAV MAL		S				Q	
LAV ELI		N			D	D	
					S	S	
730	740	750	760	770	780	790	800
LAV BRU	AGIRKVLFLD	GIDKAQDEHE	KYHSNWRAMA	SDFNLPPVVA	KEIVASCDKC	QLKGEAMHGQ	VDCSPGIWQL
ARV 2		N	E			DCTHILEGKV	I
LAV MAL	S		E		I		I
LAV ELI	Q		E	N			
810	820	830	840	850	860	870	880
LAV BRU	LVAVHVASGY	IEAEVIPAET	GOETAYFLIK	LAGRWPVKTI	HTDNGSNFTS	TTVKAACWMA	GIKQEFFGIPY
ARV 2							NPOSQGVVES
LAV MAL	I			I	VV		
LAV ELI						AA	

FIG. 3D-1

## FIG. 3D-2

890	900	910	920	930	940	950	960
LAV BRU	MNKELKKIIG	QVRDQAEHLK	TAVQMAVFIH	NFKRKGGIGG	YSAGERIVDI	IATDIQTKEI	QKQITIKIQNF
ARV 2	N	E				KK	
LAV MAL					I M		
LAV ELI				RR	I		
					I		
970	980	990	1000	1010			
LAV BRU	LMKGPALKW	KGEGAIVIQQ	NSSDIKVVPRR	KAKIIRDYGK	QMAGDDCVAS	RQDED	
ARV 2							
LAV MAL	I						
LAV ELI	I	K	V				
				G G			

ENV		SP	30	40	50	60	70	80
LAV BRU	MRVK---EKY QHLMRNGWKW	GTMLLIGILMI	CSATEKLWVT	VYYGVPMKE	ATTTLFCASD	AKAYDTEVN	VWATHACVPT	
ARV 2	K GTRRN	---	- L	M		R		
LAV MAL	REIQRN NW	---	- M	M	T	S E	I	
LAV ELI	ARGIERNQ NW K	---	- I	T	IA D	S E	A	I
					ADN			
LAV BRU	DPNPQEVVLV NVTENTNNMK	NDMVEQMHED	IIISLWDQSLK	PCVKLTLPLCV	SLKCTDL-CN	ATNTNNSNTN	SSSGEMMME-	160
ARV 2	C	N	Q		T N	- K	---	-NWKE I
LAV MAL	IE E	G	N		T N	NVN T	V	GTNACCS RTNA LK I
LAV ELI	IA E		N		T N	S E-L	RN	GTMG NV TTEEKG---
LAV BRU	KGELIKNCSEN ISTSIRGKVQ	KEYAFFYKLD	IIPIIDNDTS	-----YTLTS	INTSVITQAC	PKVSFEPIDI	HYCAPAGFAI	240
ARV 2	T D I	N L RN	VW	AS T	TNYTN R	R	T	
LAV MAL	- V	TPVGSD R	- T N	LVQ	DSDN	----S R	TD	
LAV ELI	---M	VT VIKD K	QV L R	V	SST	-NSTN R	A	
LAV BRU	LKCNKKTFNG	TGPCTNVSTV	QCTHIGIRPVV	STQLLINGSL	AAEEVVIRSA	MFTDNAKTII	VQLNQSVEIN	CTRPMTNTRK
ARV 2	K	I				D N	E A	
LAV MAL	D K	EI K				E L T N	ET T	G R
LAV ELI	RD K					AH E K T	A YQ Q	

FIG. 3E-1

# FIG. 3E-2

LAV BRU	330	340	350	360	370	380	390	400
ARV 2	Y --	W T RI	DI K	Q N E	V K	- V N	M	R
LAV MAL	G HF --	Q LY T I-V	DI R Y T N	ETE DK	Q V V	GSLL-	T	R
LAV ELI	RTP --	L Q SLY TKS-RS	TIG	Q SK Q	V R	GTLL-	I K P	T
LAV BRU	410	420	430	440	450	460	470	480
ARV 2	T N	--- -RLN	RTEG K N	I	I	C	S	T -V
LAV MAL	TSK	Q NGARL-	- S STGS	I	KT	A V	N L	NSSD
LAV ELI	TSG	NI A NNI	TES NSTNTN	Q	I K VAGR-	I	ERN L	I --
LAV BRU	490	500	510	520	530	540	550	560
ARV 2	I DT V	I	I	I	V	M	V L	A L
LAV MAL	SDN TL	R	E	I	L-	M	V	
LAV ELI	STN T	Q	E	I	L-	M		

Sequence: AFVTIGK-IG NMRQAHCNIS RAKWNATLKQ TASKLREQFG NNIKT-TIEFKQ SSGCDPEIVT HSFNCGEFF

Sequence: YCNSTQLFNS TWENSTWSTE CSNNTEGSDT ITLPCRIKQF INMMQEVGKA MYAPPISGGQI RCSSNITGLL LTRDGNN --

Sequence: NNGSEIFRPG GGDMDRNWRS ELYKVKVVKI EPLGVAPTKA KRRVVQREKR AWGI-GALFL GFLGAAGSTM GARSMTLTVQ

## FIG. 3F-1

# FIG. 3F-2

F	10	20	30	40	50	60	70	80
LAV BRU	MGGKMSKSSV	VGMPTVTERM	R----RAEPA	ADGVGAASR-	----DLEKUG	ATISSNTAAT	NAACAWLEAQ	EE-EEVGFPV
ARV 2	R M G	SAI	RAEP	V	-	D	-	-
LAV MAL	I	KI	I	---- TP T ET	V QD AVSQ	D C	AA SP N	S --- PP E
LAV ELI	I	AI	I	---- TM	V	-----	S D	SD
LAV BRU	TPQVPLRRHT	YKAAVDSLHF	LKEKGGLLEG	IHSQRQRQDIL	DLWLYUTQGY	FPPDWQNYTPC	PGVRYPLTFG	WCYKLVPVEP
ARV 2	R	L	I	W	E	I	F	F
LAV MAL	R	G F	D	VW PK	E	V	F	HS
LAV ELI	R	E L	W	W KK	E	V N	I	E
LAV BRU	DKVTEANKGE	NTSLLHPVSL	HGMDDPEREV	LEWRFDSRLA	FHHVARELHP	EYFKNC		
ARV 2	E	E	N	M	E A K	V	M	Y D
LAV MAL	EE	E	NC	I Q	E A	K K	S	Y D
LAV ELI	QE	DTE	TN	ICQ	E	Q	K N	Y D
						E K	M	FY -

A LAVbru VS.	GAG	POL	ENV			
			TOTAL	OMP	TMP	
HTLV-3 USA	512 0/0	0.8 0/0	1015 1.3	856 5/0	1.4 5/0	507 1.6
ARV-2 USA	502 12/2	3.4 12/0	1003 3.1	855 17/11	13.0 17/10	505 14.3
LAVelli ZAIRE	500 13/1	9.8 13/0	1002 5.5	853 22/14	20.7 22/14	504 25.3
LAVmal ZAIRE	505 14/7	12.0 13/0	1002 7.7	859 13/11	21.7 13/10	509 26.4
B LAVelli VS.						
LAVmal	505 1/6	10.8 0/0	1002 8.4	859 13/11	19.8 8/13	509 23.6
						350 0/1
						14.3

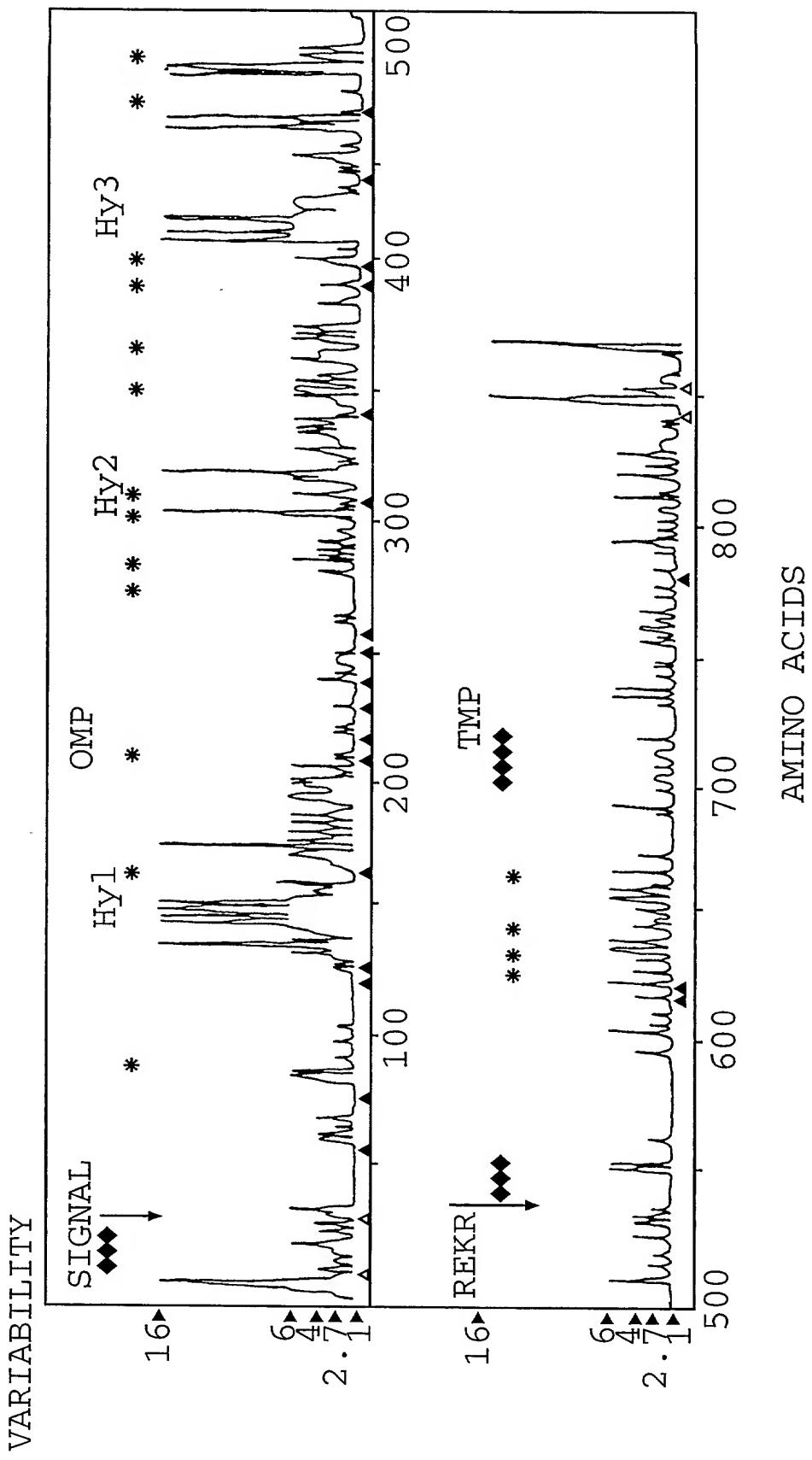
**FIG. 4A**

A LAVbru vs.	ORF F	central region			
		ORF Q	ORF R	ORF S	nd
HTLV-3 USA	206 0/0	1.5 0/0	192 0/0	0 0/1	nd 97 0/1
ARV-2 USA	210 0/4	12.6 0/0	192 0/0	10.0 0/0	9.4 11.5 0/0
LAVeli ZAIRE	206 1/1	19.4 0/0	192 0/0	10.4 0/0	96 11.5 0/0
LAVmal ZAIRE	209 2/5	27.0 0/0	192 0/0	12.6 0/0	96 10.4 0/0

B LAVeli vs.	central region			
ORF Q	ORF R	ORF S	nd	
209 3/6	22.5 0/0	192 0/0	12.0 0/0	96 6.3 0/0

**FIG. 4B**



**FIG. 5**

GAG

a

120

LAV.BRU	K	A	Q	Q	A	A	GCT	-	-	-	-	-	D	T
	AAA	GCA	CAG	CAA	GCA	GCA							GAC	ACA
ARV 2	K	A	Q	Q	A	A	GCT	A	A	-	-	-	G	T
	AAG	GCA	CAG	CAA	GCA	GCA	GCT	GCA	GCT				GGC	ACA
LAV.MAL	K	T	Q	Q	A	A	A							
	AAG	ACA	CAG	CAG	GCA	GCA	GCT	GCA	GCA	Q	Q	A	A	T
LAV.ELI	X	A	Q	Q	A	A	A	GCT	-	-	-	-	D	T
	AAG	GCA	CAG	CAA	GCA	GCA	GCT						GAC	ACA

FIG. 6A-1

b

LAV.BRU

460

G	N	F	L	Q	S	R	P	E	P	T	A	P	P	F	L	Q	S	R	P	E	P	T	A	P	P	E	E
GGG	AAT	TTT	CTT	CAG	AGC	ACA	GCA	GAG	CCA	ACA	GCC	CCA	CCA	TTT	CTT	CAG	AGC	AGA	CCA	GAG	CCA	ACA	GCC	CCA	CCA	GAA	GAG

ARV.2

470

G	N	F	L	Q	S	R	P	E	P	T	A	P	P	E	E										
GGG	AAT	TTT	CTT	CAG	AGC	ACA	GCA	GAG	CCA	ACA	GCC	CCA	CCA	-	-	-	-	-	-	-	-	-	-	GAA	GAG

LAV.MAL

G	N	F	L	Q	S	R	P	E	P	T	A	P	P	A	E									
GGG	AAT	TTC	CTT	CAG	AGC	ACA	GCA	GAG	CCA	ACA	GCC	CCA	CCA	-	-	-	-	-	-	-	-	-	GCA	GAG

LAV.ELI

G	N	F	L	Q	S	R	P	E	P	T	A	P	P	A	E									
GGG	AAC	TTT	CTC	CAA	AGC	AGA	CCA	GAG	CCA	ACA	GCC	CCA	CCA	-	-	-	-	-	-	-	-	-	GCA	GAG

**FIG. 6A-2**

c	LAV.BRU	R	M	R	20				30			
		AGA	ATG	AGA	-	-	-	-	CGA	GCT	GAG	CCA
	ARV 2	R	M	R	R	A	E	P	CGA	GCT	GAG	CCA
		AGA	ATG	AGA	CGA	GCT	GAG	CCA	CGA	GCT	GAG	CCA
	LAV.MAL	R	I	R	-	-	-	-	CGA	ACT	CCC	CCA
		AGA	ATA	AGA	-	-	-	-	R	T	P	T
		R	I	R	-	-	-	-	AGA	ACT	AAT	CCA
	LAV.ELI	AGA	ATA	AGA					GCA			
d	LAV.BRU	V	G	A	A	S	R					D
		GTG	GGA	GCA	GCA	TCT	CGA	-	-	-	-	- GAC
	ARV 2	V	G	A	V	A	R					D
		GTG	GGA	GCA	GTA	TCT	CGA	-	-	-	-	- GAC
		V	G	A	V	S	R	D	A	V	S	Q
		GTA	GGA	GCA	GTA	TCT	CAA	GAT	GCA	GTA	TCT	CAA
	LAV.MAL	V	G	A	V	S	R	-	-	-	-	D
		GTA	GGA	GCA	GTA	TCT	CGA	-	-	-	-	- GAC
	LAV.ELI											

**FIG. 6A-3**

old people have not yet been able to get out of the mud and those in their houses have not yet been able to get home.

		ENV 20											
		L						T					
		W	R	W	G	W	K	W	G	W	M	W	G
LAV.BRU	CAG	CAC	TTG	TGG	ACA	TGG	GGC	TGG	AAA	TGG	GGC	ACC	ATG
ARV 2	CAG	CAC	TTG	TGG	AGA	TGG	GGC	-	-	-	-	ACC	TTG
LAV.MAL	CAA	AAC	TGG	TGG	AGA	TGG	GGC	-	-	-	-	ATG	ATG
LAV.ELI	CAA	AAC	TGG	TGG	AAA	TGG	GGC	-	-	-	-	ATC	ATG

LAV.BRU	140	150
L <sub>TA</sub> K <sub>AAG</sub> C <sub>TGC</sub> T <sub>ACT</sub> D <sub>GAT</sub> L <sub>TTG</sub> - G <sub>GGG</sub> N <sub>AAT</sub> A <sub>GCT</sub> T <sub>ACT</sub> N <sub>AAT</sub> T <sub>ACC</sub> N <sub>AAT</sub> T <sub>ACC</sub> S <sub>AAT</sub> T <sub>ACC</sub> N <sub>AAT</sub> S <sub>AGT</sub> A <sub>AGT</sub> AG <sub>T</sub>		
M <sub>ATG</sub> M <sub>ATG</sub> M <sub>ATG</sub> E <sub>GAG</sub> - K <sub>AAA</sub> G <sub>CAG</sub> E <sub>GAG</sub> I <sub>ATA</sub>		
ARG 2		
L <sub>TTA</sub> N <sub>AAT</sub> C <sub>TGC</sub> T <sub>ACT</sub> D <sub>GAT</sub> L <sub>TTG</sub> - G <sub>GGG</sub> K <sub>AAG</sub> A <sub>GCT</sub> T <sub>ACT</sub> N <sub>AAT</sub> T <sub>ACC</sub> N <sub>AAT</sub> S <sub>AGT</sub> - - - - -		
W <sub>TTG</sub> K <sub>AAA</sub> E <sub>GAA</sub> E <sub>GAA</sub> E <sub>GAA</sub> I <sub>ATA</sub> K <sub>AAA</sub> G <sub>GAA</sub> G <sub>GAA</sub> G <sub>GAA</sub> I <sub>ATA</sub>		

FIG. 6B-1

## LAV.MAL

L	N	C	T	N	V	N	G	T	A	V	N	G	T	N	A	G	S	N	R	T	N	A	E
TTA	AAC	TGC	ACT	AAT	GTG	AAT	GGG	ACT	GCT	GTG	AAT	GGG	ACT	AAT	GCT	GGG	AGT	AAT	AGG	ACT	AAT	GCA	GAA
L	K	M	E	I		G	E	V															
TTG	AAA	ATG	GAA	ATT	-	GGA	GAA	GTG															

## LAV.ELI

L	N	C	S	D	E	-	L	R	N	N	G	T	M	G	N	N	V	T	T	E	E	K	
TTA	AAC	TGT	AGT	GAT	GAA	-	TTG	AGG	AAC	AAT	GGC	ACT	ATG	GGG	AAC	AAT	GTC	ACT	ACA	GAG	GAG	AAA	
G							M						ATG										
GGA	-	-	-	-	-	-		-	-	-	-	-											

**FIG. 6B-2**

## FIG. 6B-3

BRU. LAV. ARV. LAV. LAV. LAV. ARV.

g D N D T T S 200  
LAV. BRU GAT AAT GAT ACT ACC AGC - - - - - Y T L  
ARV 2 D N A S T T T N Y T N Y R  
GAT AAT GCT AGT ACT ACT ACC AAC TAT ACC AAC TAT AGG TTG

LAV. MAL D D S D N S S 410  
GAT GAT AGT GAT MAT AGT AGT - - - - - Y R L  
D N D S S T N Y 420  
GAC AAT GAT AGT AGT ACC - AAT AGT ACC AAT TAT AGG TTA

LAV. ELI D N D S S T N Y 430  
GAC AAT GAT AGT AGT ACC - AAT AGT ACC AAT TAT AGG TTA

h

LAV. BRU F N S T W F N S T W S T E G S N N T E G  
TGT AAT TCA ACA CAA CTG TTT AAT AGT ACT TGG TTT AAT AGT ACT TGG AGT ACT GAA GGG TCA AAT AAC ACT GAA GGA

S D T I  
AGT GAC ACA ATC

ARV 2

C N T Q L F N N T W R L N H T K G T K G  
TGT AAT ACA ACA CAA CTG TTT AAT AGT ACA TGG - - - - AGG TTA AAT CAC ACT GAA GGA ACT AAA GGA

N D T I  
AAT GAC ACA ATC

...TGT AAT ACA TCA AAA CTG TTT AAT AGT ACA TGG CAG AAT AAT GGT GCA AGA CTA - - AGT AAT ACA GAG TCA

LAV.MAL

C N T S K L F N S T W Q N N G A R L S N S T E S  
TGT AAT ACA TCA AAA CTG TTT AAT AGT ACA TGG CAG AAT AAT GGT GCA AGA CTA - - AGT AAT ACA GAG TCA

T G S I  
ACT GGT AGT ATC

LAV.ELI

C N T S G L F N S T W N I S A W N N I T E S N N S T  
TGT AAT ACA TCA GGA CTG TTT AAT AGT ACA TGG AAT ATT AGT GCA TGG AAT ATT ACA GAG TCA ATT ACA GAG TCA ATT AGC ACA

N T N I  
AAC ACA AAC ATC

**FIG. 6B-4**

LAV.MAL

→R

GGTCTCTTTGTTAGACCAGGTCGAGCCGGAGCTCTGGCTAGCAAGGAACCCACTG  
CTTAAGCCTCAATAAAGCTTGCCTTGAGTGCCTCAAGCAGTGTGTGCCATCTGTTGTGT  
100  
GACTCTGGTAACTAGAGATCCCTCAGACCACTCTAGACGGTGTAAAAATCTCTAGCAGTG  
GCGCCGAACAGGGACTTTAAAGTGAAGTAACAGGGACTCGAAAGCGGAAGTTCCAGAG  
200  
AAGTTCTCTCGACGCAGGACTCGGCTTGAGGTGCACACAGCAAGAGGCCAGAGCGGC  
300  
GAG MetGlyAlaArg  
GAATGGTGAGTACGCCATTGGTACTAGCGGAGGCTAGAAGGAGAGAGATGGGTGCGAG  
AlaSerValLeuSerGlyGlyLysLeuAspAlaTrpGluLysIleArgLeuArgProGly  
AGCGTCAGTATTAAGCGGGGGAAAATTAGATGCATGGGAGAAAATCGGTTAAGGCCAGG  
400  
GlyLysLysLysTyrArgLeuLysHisLeuValTrpAlaSerArgGluLeuGluArgPhe  
GGGAAAGAAAAATATAGACTGAAACATTAGTATGGGCAAGCAGGGAGCTGGAAAGATT  
AlaLeuAsnProGlyLeuLeuGluThrGlyGluGlyCysGlnGlnIleMetGluGlnLeu  
CGCACTTAACCCCTGGCCTTTAGAACAGGAGAAGGATGTCAACAAATAATGGAACAGCT  
500  
GlnSerThrLeuLysThrGlySerGluGluIleLysSerLeuTyrAsnThrValAlaThr  
ACAATCAAACCTCAAGACAGGATCAGAACAAATTAAATCATTATATAATACAGTAGCAAC  
600  
LeuTyrCysValHisGlnArgIleAspValLysAspThrLysGluAlaLeuAspLysIle  
CCTCTATTGTGTACATCAAAGGATAGATGTAAAAGACACCAAGGAAGCGCTAGATAAAAT  
700  
GluGluIleGlnAsnLysSerArgGlnLysThrGlnGlnAlaAlaAlaAlaGlnGlnAla  
AGAGGAAATACAAATAAGAGCAGGCAAAAGACACAGCAGGCAGCAGCTGCACAGCAGGC  
AlaAlaAlaThrLysAsnSerSerValSerGlnAsnTyrProIleValGlnAsnAla  
800  
AGCAGCTGCCACAAAAACAGCAGCAGTCAGTCAAAATTACCCATAGTGCACAAATGC  
GlnGlyGlnMetIleHisGlnAlaIleSerProArgThrLeuAsnAlaTrpValLysVal  
ACAAGGGCAAATGATACATCAGGCCATATCACCTAGGACTTGAATGCATGGTGAAAGT  
900  
IleGluGluLysAlaPheSerProGluValIleProMetPheSerAlaLeuSerGluGly  
AATAGAAGAAAAGGCTTCAGCCCAGAAGTGTACCCATGTTCTCAGCATATCAGAGGG  
AlaThrProGlnAspLeuAsnMetMetLeuAsnIleValGlyGlyHisGlnAlaAlaMet  
GGCCACCCCCACAAGATTAAATATGATGCTGAACATAGTTGGAGGACACCAGGCAGCTAT  
GlnMetLeuLysAspThrIleAsnGluGluAlaAlaAspTrpAspArgValHisProVal  
GCAAATGTTAAAAGATACCATCAATGAGGAAGCTGCAGACTGGACAGGGTACATCCAGT  
1000  
HisAlaGlyProIleProProGlyGlnMetArgGluProArgGlySerAspIleAlaGly  
ACATGCAGGGCCTATTCCCCCAGGCCAGATGAGAGAACCAAGAGGAAGTGACATAGCAGG

**FIG. 7A**

Thr Thr Ser Thr Leu Gln Glu Gln Ile Gly Trp Met Thr Ser Asn Pro Pro Ile Pro Val  
 AACTACTAGTACCCTCAAGAACAAATAGGATGGATGACAAGCAACCCACCTATCCCAGT  
 1100  
 Gly Asp Ile Tyr Lys Arg Trp Ile Ile Leu Gly Leu Asn Lys Ile Val Arg Met Tyr Ser  
 GGGAGAACATCTATAAAAGATGGATAATCCTGGGATTAAATAAAATAGTAAGAATGTATAG  
 1200  
 Pro Val Ser Ile Leu Asp Ile Arg Gln Gly Pro Lys Glu Pro Phe Arg Asp Tyr Val Asp  
 CCCTGTCAGCATTGGACATAAGACAAGGGCAAAGGAACCTTTAGAGACTATGTAGA  
 1300  
 Arg Phe Phe Lys Thr Leu Arg Ala Glu Gln Ala Thr Gln Glu Val Lys Asn Trp Met Thr  
 TAGGTTCTTAAACTCTCAGAGCTGAGCAAGCTACACAGGAGGTAAAAATTGGATGAC  
 1400  
 Glu Thr Leu Leu Val Gln Asn Ala Asn Pro Asp Cys Lys Thr Ile Leu Lys Ala Leu Gly  
 AGAACCTTGCTGGTCCAAATGCGAATCCAGACTGTAAGACCATTAAAGCATTAGG  
 1500  
 Pro Gly Ala Thr Leu Glu Glu Met Met Thr Ala Cys Gln Gly Val Gly Gly Pro Ser His  
 ACCAGGGCTACATTAGAAGAAATGATGACAGCATGCCAGGGAGTGGGAGGACCCAGTCA  
 1600  
 Lys Ala Arg Val Leu Ala Glu Ala Met Ser Gln Ala Thr Asn Ser Thr Ala Ala Ile Met  
 TAAAGCAAGAGTTTGCTGAGGCAATGAGCCAAGCAACAAATTCAACTGCTGCCATAAT  
 1700  
 Met Gln Arg Gly Asn Phe Lys Gly Gln Lys Arg Ile Lys Cys Phe Asn Cys Gly Lys Glu  
 GATGCAGAGAGGTAAATTAAAGGGCCAGAAAAGAATTAAAGTGTTCAACTGTGGCAAAGA  
 1800  
 Gly His Leu Ala Arg Asn Cys Arg Ala Pro Arg Lys Lys Gly Cys Trp Lys Cys Gly Lys  
 AGGACACCTAGCCAGAAATTGCAGGGCCCCTAGGAAAAAGGGCTGTTGGAAATGTGGGAA  
 1900  
 →POL  
 Phe Phe Arg Glu Asn Leu  
 Glu Gly His Gln Met Lys Asp Cys Thr Glu Arg Gln Ala Asn Phe Leu Gly Lys Ile Trp  
 GGAAGGACACCAAATGAAAGACTGCACTGAGAGACAGGCTAAATTAGGGAAAATTG  
 2000  
 Ala Phe Pro Gln Gly Lys Ala Arg Glu Phe Pro Ser Glu Gln Thr Arg Ala Asn Ser Pro  
 Pro Ser His Lys Gly Arg Pro Gly Asn Phe Leu Gln Ser Arg Pro Glu Pro Thr Ala Pro  
 GCCTTCCCACAAGGGAAAGGCCAGGGATTTCCTTCAGAGCAGACCAGGCCAACAGCCCC  
 2100  
 Thr Ser Arg Glu Leu Arg Val Trp Gly Gly Asp Lys Thr Leu Ser Glu Thr Gly Ala Glu  
 Pro Ala Glu Ser Phe Gly Phe Gly Glu Ile Lys Pro Ser Gln Lys Gln Glu Gln Lys  
 ACCAGCAGAGAGCTCGGGTTGGGGAGGAGATAAAACCCCTCTCAGAACAGGAGCAGAA  
 2200  
 Arg Gln Gly Ile Val Ser Phe Ser Phe Pro Gln Ile Thr Leu Trp Gln Arg Pro Val Val  
 Asp Lys Glu Leu Tyr Pro Leu Ala Ser Leu Lys Ser Leu Phe Gly Asn Asp Gln Leu Ser  
 AGACAAGGAATTGTATCCTTAGCTCCCTCAAATCACTCTTGCAACGACCAGTTGTC  
 GAG ←  
 Thr Val Arg Val Gly Gly Gln Leu Lys Glu Ala Leu Leu Asp Thr Gly Ala Asp Asp Thr  
 Gln  
 ACAGTAAGAGTAGGAGGACAGCTAAAGAACAGCTCTATTAGACACAGGAGCAGATGATACA  
 2300  
 Val Leu Glu Glu Ile Asn Leu Pro Gly Lys Trp Lys Pro Lys Met Ile Gly Gly Ile Gly  
 GTATTAGAAGAAATAATTGCCAGGAAAATGGAACCAAAATGATAGGGGGATTGGA  
 2400  
 Gly Phe Ile Lys Val Arg Gln Tyr Asp Gln Ile Leu Ile Glu Ile Cys Gly Lys Lys Ala  
 GGTTTATCAAAGTAAGACAGTATGATCAAATACTTATAGAAATTGTGGAAAAAGGCT  
 2500

**FIG. 7B**

IleGlyThrIleLeuValGlyProThrProValAsnIleIleGlyArgAsnMetLeuThr  
 ATAGGTACAATATTGGTAGGACCTACACCTGTCAACATAATTGGACGAAATATGTTGACT  
 2100  
 GlnIleGlyCysThrLeuAsnPheProIleSerProIleGluThrValProValLysLeu  
 CAGATTGGTTGTACTTAAATTTCACATTAGTCCTATTGAGACTGTACCAGTAAAATTA  
 LysProGlyMetAspGlyProArgValLysGlnTrpProLeuThrGluGluLysIleLys  
 AAGCCAGGGATGGATGGCCCAAGGGTTAACAAATGCCATTGACAGAAGAAAAATAAAA  
 2200  
 AlaLeuThrGluIleCysLysAspMetGluLysGluGlyLysIleLeuLysIleGlyPro  
 GCATTAACAGAAATTGTAAAGATATGGAAAAGGAAGGAAAAATTAAATTGGGCCT  
 GluAsnProTyrAsnThrProValPheAlaIleLysLysLysAspSerThrLysTrpArg  
 GAAAATCCATACAATACTCCAGTATTGCCATAAAAGAAAAAAGACAGCACTAAATGGAGA  
 2300  
 LysLeuValAsnPheArgGluLeuAsnLysArgThrGlnAspPheTrpGluValGlnLeu  
 AAATTAGTGAATTTCAGAGAGCTTAATAAAAGAACTCAAGATTGGAAAGTTCAATTAA  
 2400  
 GlyIleProHisProAlaGlyLeuLysLysLysSerValThrValLeuAspValGly  
 GGAAATACCACATCCTGCTGGGTTGAAAAAGAAAAATCAGTCACAGTATTGGATGTGGGG  
 AspAlaTyrPheSerValProLeuAspGluAspPheArgLysTyrThrAlaPheThrIle  
 GATGCATATTTCAGTCCCTTAGATGAAGATTTCAGGAAGTATACTGCATTCACTATA  
 2500  
 ProSerIleAsnAsnGluThrProGlyIleArgTyrGlnTyrAsnValLeuProGlnGly  
 CCCAGTATTAAATAATGAGACACCAGGGATTAGATATCAGTACAATGTGCTACCACAGGG  
 TrpLysGlySerProAlaIlePheGlnSerSerMetThrLysIleLeuGluProPheArg  
 TGGAAAGGATCACCAAGCAATTCCAGAGTAGCATGACAAAAATCTTAGAACCCCTTT AGA  
 2600  
 ThrLysAsnProGluIleValIleTyrGlnTyrMetAspAspLeuTyrValGlySerAsp  
 ACAAAAATCCAGAAATAGTCATATACCAATACATGGATTTGTATGTAGGGTCTGAT  
 2700  
 LeuGluIleGlyGlnHisArgThrLysIleGluGluLeuArgGluHisLeuLeuLysTrp  
 TTAGAAATAGGACAACATAGAACAAAATAGAGGAACTAAGAGAACATCTATTGAAATGG  
 GlyPheThrThrProAspLysLysHisGlnLysGluProProPheLeuTrpMetGlyTyr  
 GGATTACCAACCAGAACAAAAGCATCAGAAAGAACCCCCATTCTTGGATGGGTAT  
 2800  
 GluLeuHisProAspLysTrpThrValGlnProIleGlnLeuProAspLysGluSerTrp  
 GAACTCCACCCTGACAATGGACAGTGCAGCCTATACAACGCCAGACAAGGAAAGCTGG  
 ThrValAsnAspIleGlnLysLeuValGlyLysLeuAsnTrpAlaSerGlnIleTyrPro  
 ACTGTCAATGATATACAGAAATTGGTGGGAAAACCAAATTGGCAAGTCAGATTATCCA  
 2900  
 GlyIleLysValLysGlnLeuCysLysLeuLeuArgGlyAlaLysAlaLeuThrAspIle  
 GGAAATTAAAGTAAAGCAATTATGTAACCTCCTTAGGGAGCAAAAGCACTAACAGACATA  
 3000  
 ValProLeuThrAlaGluAlaGluLeuGluLeuAlaGluAsnArgGluIleLeuLysGlu  
 GTACCATTAACTGCAGAGGCAGAAATTAGAACATTGGCAGAGAACAGGGAAATTCTAAAAGAA

**FIG. 7C**

ProValHisGlyValTyrTyrAspProSerLysAspLeuIleAlaGluIleGlnLysGln  
CCAGTGCATGGGGTATATTGACCCATCAAAGACTTAATAGCAGAAATACAGAAGCAG  
3100  
GlyGlnGlyGlnTrpThrTyrGlnIleTyrGlnGlnTyrLysAsnLeuLysThrGly  
GGGCAAGGTCAATGGACATATCAAATATAACCAAGAGCAATATAAAATCTGAAAACAGGG  
LysTyrAlaArgIleLysSerAlaHisThrAsnAspValLysGlnLeuThrGluAlaVal  
AAGTATGCAAGAATAAAGTCTGCCACACTAATGATGTAAAACAATTACAGAAGCAGTG  
3200  
GlnLysIleAlaGlnGluSerIleValIleTrpGlyLysThrProLysPheArgLeuPro  
CAAAAGATAGCCCAAGAAAGCATAGTAATATGGGGAAAAACTCCTAAATTAGACTACCC  
3300  
IleGlnLysGluThrTrpGluAlaTrpTrpThrGluTyrTrpGlnAlaThrTrpIlePro  
ATACAAAAAGAACATGGGAGGCATGGTGGACAGAATATTGGCAAGCCACCTGGATCCCT  
GluTrpGluPheValAsnThrProProLeuValLysLeuTrpTyrGlnLeuGluThrGlu  
GAATGGGAGTTGTCAATACTCCTCCCTAGTAAACTATGGTACCAAGTTAGAAACAGAA  
3400  
ProIleValGlyAlaGluThrPheTyrValAspGlyAlaAlaAsnArgGluThrLysLys  
CCCATAGTAGGAGCAGAAACTTCTATGTAGATGGGCAGCTAATAGAGAAACTAAAAAG  
GlyLysAlaGlyTyrValThrAspArgGlyArgGlnLysValValSerLeuThrGluThr  
GGAAAAGCAGGATATGTACTGACAGAGGAAAGACAAAAGGTTGTCTCCTTAAGTCAA  
3500  
ThrAsnGlnLysThrGluLeuGlnAlaIleHisLeuAlaLeuGlnAspSerGlySerGlu  
ACAAATCAGAAGACTGAATTACAAGCAATCCACTTAGCTTACAGGATTAGGATCAGAA  
3600  
ValAsnIleValThrAspSerGlnTyrAlaLeuGlyIleIleGlnAlaGlnProAspLys  
GTAAACATAGTAACAGACTCACAGTATGCATTAGGGATTATTCAAGCACAACCAGATAAA  
SerGluSerGluIleValAsnGlnIleIleGluGlnLeuIleGlnLysAspLysValTyr  
AGTGAATCAGAGATTGTTAATCAAATAATAGAGCAATTACAGAAGGACAAGGTCTAC  
3700  
LeuSerTrpValProAlaHisLysGlyIleGlyGlyAsnGluGlnValAspLysLeuVal  
CTGTCATGGTACCAAGCACACAAAGGGATTGGAGGAAATGAACAGTAGATAAAATTAGTC  
SerSerGlyIleArgLysValLeuPheLeuAspGlyIleAspLysAlaGlnGluGluHis  
AGCAGTGGAAATCAGAAAGGTACTATTTTAGATGGGATAGATAAGGCTCAAGAAGAACAT  
3800  
GluLysTyrHisSerAsnTrpArgAlaMetAlaSerAspPheAsnLeuProProIleVal  
GAAAAATATCACAGCAATTGGAGAGCAATGGCTAGTGACTTAATCTACCACCTATAGTA  
3900  
AlaLysGluIleValAlaSerCysAspLysCysGlnLeuLysGlyGluAlaMetHisGly  
GCGAAGGAAATAGTAGCCAGCTGTGATAACTAAAGGGAAAGCCATGCATGGAA  
GlnValAspCysSerProGlyIleTrpGlnLeuAspCysThrHisLeuGluGlyLysIle  
CAAGTAGACTGTAGTCCAGGGATATGGCAATTAGATTGCACACATCTAGAAGGAAAAATA  
4000  
IleIleValAlaValHisValAlaSerGlyTyrIleGluAlaGluValIleProAlaGlu  
ATCATAGTAGCAGTCCATGTAGCCAGTGGATATAGAAGCAGAAGTTATCCCAGCAGAA  
ThrGlyGlnGluThrAlaTyrPheIleLeuLysLeuAlaGlyArgTrpProValLysVal  
ACAGGACAGGAGACAGCATACTTATACTAAAGTAGCAGGAAGATGCCAGTAAAAGTA  
4100

**FIG. 7D**

**FIG. 7E**

ProGlnArgGluProHisAsnGluTrpThrLeuGluLeuLeuGluLeuLysGlnGlu  
 HisArgGlySerHisThrMetAsnGlyHis  
 CCACAGAGGGAGGCCACACAATGAATGGACAT TAGAACTTTAGAGGAGCTTAAGCAAGAA  
 5200

AlaValArgHisPheProArgIleTrpLeuHisSerLeuGlyGlnHisIleTyrGluThr  
 GCTGTCAGACACTTCCTAGGATATGGCTCCATAGTTAGGACAACATATCTATGAAACT

TyrGlyAspThrTrpGluGlyValGluAlaIleIleArgSerLeuGlnGlnLeuLeuPhe  
 TATGGGGATACCTGGGAAGGAGTTGAAGCTATAATAAGAAGTCTGCAACAACTGCTGTT  
 5300

IleHisPheArgIleGlyCysGlnHisSerArgIleGlyIleThrArgGlnArgArgAla  
 ATTCACTTCAGAATTGGGTGTCAACATAGCAGAATAGGCATTACTCGACAGAGAAGAGCA  
 5400

ArgAsnGlySerSerArgSer  
 MetAspProValAspProAsnLeuGluProTrpAsnHisProGlySerGlnProArg  
 AGAAATGGATCCAGTAGATCCTAACTTAGAGGCCCTGGAACCATCCAGGGAGTCAGCCTAG

ThrProCysAsnLysCysTyrCysLysLysCysCysTyrHisCysGlnMetCysPheIle  
 GACGCCTTGTAAATAAGTGTATTGTAAAAAGTGTGCTATCATTGCCAAATGTGCTTCAT  
 5500

ThrLysGlyLeuGlyIleSerTyrGlyArgLysLysArgArgGlnArgArgArgProPro  
 AACGAAAGGCTTAGGCATCTCCTATGGCAGGAAGAACGGAGACAGCGACGAAGACCTCC

S ←  
 GlnGlyAsnGlnAlaHisGlnAspProLeuProGluGln  
 TCAGGGCAATCAGGCTCATCAAGATCCTCTACCAAGAGCAGTAAGTAGTATATGTAATACA  
 5600

ACCTTAGTGTATTAGCAATAGTAGCATTAGTAGTAACGCTAATAATAGCAATAGTGT  
 5700

GTGGACCATAGTATTAGAAATTAGGAAAATAAGAACAAAGGAAAATAGACAGGTT  
 ENV

MetArgValArgGluIleGlnArg  
 GATTGATAGAATAAGAGAAAGAGCAGAACAGATAGTGGCAATGAGAGTGAGGGAGATACAGA  
 5800

AsnTyrGlnAsnTrpTrpArgTrpGlyMetMetLeuLeuGlyMetLeuMetThrCysSer  
 GGAATTATCAAACACTGGTGGAGATGGGGCATGATGCTCCTGGATGTTGATGACCTGTA

IleAlaGluAspLeuTrpValThrValTyrTyrGlyValProValTrpLysGluAlaThr  
 GTATTGCAGAACAGATTGTGGGTACAGTTATTATGGGGTACCTGTGTGGAAAGAACCAA  
 5900

ThrThrLeuPheCysAlaSerAspAlaLysSerTyrGluThrGluValHisAsnIleTrp  
 CCACTACTCTATTGTGCATCAGATGCTAAATCATATGAAACAGAACATACATAACATCT  
 6000

AlaThrHisAlaCysValProThrAspProAsnProGlnGluIleGluLeuGluAsnVal  
 GGGCTACACATGCCTGTGTACCCACGGACCCACAAGAAATAGAACACTGGAAAATG

ThrGluGlyPheAsnMetTrpLysAsnAsnMetValGluGlnMetHisGluAspIleIle  
 TCACAGAAGGGTTAACATGTGGAAAAATAACATGGTGGAGCAGATGCATGAGGATATAA  
 6100

**FIG. 7F**

SerLeuTrpAspGlnSerLeuLysProCysValLysLeuThrProLeuCysValThrLeu  
 TCAGTTTATGGATCAAAGCCTAAAACCATGTGTAAAGCTAACCCACTCTGTGTCACCT

AsnCysThrAsnValAsnGlyThrAlaValAsnGlyThrAsnAlaGlySerAsnArgThr  
 TAAACTGCACTAATGTGAATGGACTGCTGTGAATGGACTAATGCTGGAGTAATAGGA  
 6200

AsnAlaGluLeuLysMetGluIleGlyGluValLysAsnCysSerPheAsnIleThrPro  
 CTAATGCAGAATTGAAAATGGAAATTGGAGAAGTGAAGAACTGCTCTTCAATATAACCC  
 6300

ValGlySerAspLysArgGlnGluTyrAlaThrPheTyrAsnLeuAspLeuValGlnIle  
 CAGTAGGAAAGTGATAAAAGGCAAGAATATGCAACTTTTATAACCTTGATCTAGTACAAA

AspAspSerAspAsnSerSerTyrArgLeuIleAsnCysAsnThrSerValIleThrGln  
 TAGATGATAGTGATAATAGTAGTTAGGCTAATAATTGTAATACCTCAGTAATTACAC  
 6400

AlaCysProLysValThrPheAspProIleProIleHisTyrCysAlaProAlaGlyPhe  
 AGGCTTGTCCAAAGGTAACCTTGATCCAATTCCCACATATTGTGCCAGCTGGTT

AlaIleLeuLysCysAsnAspLysLysPheAsnGlyThrGluIleCysLysAsnValSer  
 TTGCAATTCTAAAGTGTAAATGATAAGAAGTTCAATGGAACGGAAATATGTAAAATGTCA  
 6500

ThrValGlnCysThrHisGlyIleLysProValValSerThrGlnLeuLeuAsnGly  
 GTACAGTACAATGTACACATGGAATTAGCCAGTGGTGTCAACTCAACTGCTGTTAAATG  
 6600

SerLeuAlaGluGluGluIleMetIleArgSerGluAsnLeuThrAspAsnThrLysAsn  
 GCAGTCTAGCAGAAGAAGAGATAATGATTAGATCTGAAAATCTCACAGACAATACTAAAA

IleIleValGlnLeuAsnGluThrValThrIleAsnCysThrArgProGlyAsnAsnThr  
 ACATAATAGTACAGCTTAATGAAACTGTAACAATTAAATTGTACAAGGCCTGGAAACAATA  
 6700

ArgArgGlyIleHisPheGlyProGlyGlnAlaLeuTyrThrGlyIleValGlyAsp  
 CAAGAAGAGGGATACATTGGCCAGGGCAAGCACTCTACAAACAGGGATAGTAGGAG

IleArgArgAlaTyrCysThrIleAsnGluThrGluTrpAspLysThrLeuGlnGlnVal  
 ATATAAGAAGAGCATATTGTACTATTAAATGAAACAGAATGGATAAAACTTACAACAGG  
 6800

AlaValLysLeuGlySerLeuLeuAsnLysThrLysIleIlePheAsnSerSerGly  
 TAGCTGTAAAATAGGAAGCCTCTAACAAAACAAAATAATTGTTAATTCATCCTCAG  
 6900

GlyAspProGluIleThrThrHisSerPheAsnCysArgGlyGluPhePheTyrCysAsn  
 GAGGGGACCCAGAAATTACAACACACAGTTAATTGTAGAGGGGAATTTCTACTGTA

ThrSerLysLeuPheAsnSerThrTrpGlnAsnAsnGlyAlaArgLeuSerAsnSerThr  
 ATACATCAAACACTGTTAATAGTACATGGCAGAATAATGGTGCAAGACTAAGTAATAGCA  
 7000

GluSerThrGlySerIleThrLeuProCysArgIleLysGlnIleIleAsnMetTrpGln  
 CAGAGTCAACTGGTAGTATCACACTCCCATGCAGAATAAAACAAATTATAAATATGTGGC

LysThrGlyLysAlaMetTyrAlaProProIleAlaGlyValIleAsnCysLeuSerAsn  
 AGAAAACAGGAAAAGCTATGTATGCCCTCCCATCGCAGGAGTCATCAACTGTTATCAA  
 7100

IleThrGlyLeuIleLeuThrArgAspGlyGlyAsnSerSerAspAsnSerAspAsnGlu  
 ATATTACAGGGCTGATATTAAACAAGAGATGGTGGAAATAGTAGTGACAATAGTGACAATG  
 7200

**FIG. 7G**

ThrLeuArgProGlyGlyGlyAspMetArgAspAsnTrpIleSerGluLeuTyrLysTyr  
 AGACCTTAAGACCTGGAGGAGATATGAGGGACAATTGGATAAGTGAATTATATAAAT

LysValValArgIleGluProLeuGlyValAlaProThrLysAlaLysArgArgValVal  
 ATAAAGTAGTAAGAATTGAACCCCTAGGAGTAGCACCCACCAAGGGCAAAGAGAAGAGTGG  
 7300

GluArgGluLysArgAlaIleGlyLeuGlyAlaMetPheLeuGlyPheLeuGlyAlaAla  
 TGGAAAGAGAAAAAGAGCAATAGGACTAGGAGCCATGTTCCCTGGGTCTGGGAGCAG

GlySerThrMetGlyAlaAlaSerLeuThrLeuThrValGlnAlaArgGlnLeuLeuSer  
 CAGGAAGCACGATGGCGCAGCGTCACTAACGCTGACGGTACAGGCCAGACAGTTACTGT  
 7400

GlyIleValGlnGlnGlnAsnAsnLeuLeuArgAlaIleGluAlaGlnGlnHisLeuLeu  
 CTGGTATAGTGCAACAGAAAACAATTGCTGAGGGCTATAGAGGCGAACAGCATCTGT  
 7500

GlnLeuThrValTrpGlyIleLysGlnLeuGlnAlaArgValLeuAlaValGluArgTyr  
 TGCAACTCACGGTCTGGGCATTAACAGCTCCAGGCAAGAGTCCTGGCTGTGGAAAGAT

LeuGlnAspGlnArgLeuLeuGlyMetTrpGlyCysSerGlyLysHisIleCysThrThr  
 ACCTACAGGATCAACGGCTCCTAGGAATGTGGGGTTGCTCTGGAAAACACATTGACCCA  
 7600

PheValProTrpAsnSerSerTrpSerAsnArgSerLeuAspAspIleTrpAsnAsnMet  
 CATTGTGCCTTGGAACTCTAGTTGGAGTAATAGATCTCTAGATGACATTGGAATAATA

ThrTrpMetGlnTrpGluLysGluIleSerAsnTyrThrGlyIleIleTyrAsnLeuIle  
 TGACCTGGATGCAGTGGAAAAGAAATTAGCAATTACACAGGCATAATATAACACTTAA  
 7700

GluGluSerGlnIleGlnGlnGluLysAsnGluLysGluLeuLeuGluLeuAspLysTrp  
 TTGAAGAATCGCAAATCCAGCAAGAAAAGAATTATTGGAATTGGACAAAGT  
 7800

AlaSerLeuTrpAsnTrpPheSerIleSerLysTrpLeuTrpTyrIleArgIlePheIle  
 GGGCAAGTTGTGGATTGGTTAGCATATCAAATGGCTGTGGTATATAAGAATATTCA

IleValValGlyGlyLeuIleGlyLeuArgIleIlePheAlaValLeuSerLeuValAsn  
 TAATAGTAGTAGGAGGCTTAATAGGTTAAGAATAATTGGCTGTGCTTCTTAGTAA  
 7900

ArgValArgGlnGlyTyrSerProLeuSerLeuGlnThrLeuLeuProThrProArgGly  
 ATAGAGTTAGGCAGGGATACTCACCTCTGTCGGCAGACCCCTCCAACACCGAGGG

ProProAspArgProGluGlyIleGluGluGluGlyGlyGluGlnGlyArgGlyArgSer  
 GACCACCCGACAGGCCGAAGGAATAGAAGAAGAAGGTGGAGAGCAAGGCAGAGGCAGAT  
 8000

IleArgLeuValAsnGlyPheSerAlaLeuIleTrpAspAspLeuArgAsnLeuCysLeu  
 CAATTGATTGGTGAACGGATTCTCAGCACTTATCTGGACGACCTGAGGAACCTGTGCC  
 8100

PheSerTyrHisArgLeuArgAspLeuLeuLeuIleAlaThrArgIleValGluLeuLeu  
 TCTTCAGTTACCAACCGCTTGAGAGACTTACTCTTAATTGCAACGAGGATTGTGGAACTTC

GlyArgArgGlyTrpGluAlaLeuLysTyrLeuTrpAsnLeuLeuGlnTyrTrpGlyGln  
 TGGGACGCAGGGGGTGGGAAGCCCTCAAATATCTGTGGAATCTCCTGCAATATTGGGTC  
 8200

**FIG. 7H**